TICAM Report 02-34

Analysis of a subdomain-based error estimator for finite element approximations of elliptic problems

S. Prudhomme, F. Nobile, L. Chamoin, and J.T. Oden

Texas Institute for Computational and Applied Mathematics
The University of Texas at Austin
Austin, Texas 78712

Abstract

In this paper we analyse a sub-domain residual error estimator for finite element approximations of elliptic problems. It is obtained by solving local problems on patches of elements in weighted spaces and provides for an upper bound on the energy norm of the error when the local problems are solved in sufficiently enriched discrete spaces. A guaranteed lower bound on the error is also derived by a simple postprocess of the solutions to the local problems. Numerical tests show very good effectivity indices for both the upper and lower bounds and a strong reliability of this estimator even for coarse meshes.

1 Introduction

The use of residual-based methods for *a posteriori* estimation of errors in finite element approximations of elliptic boundary-value problems has become important in applications as methods for assessing quality of numerical solutions and providing a basis for adaptive meshing. A survey of such methods can be found in the monograph of Ainsworth and Oden [1]. Among early residual methods was the subdomain-residual method of Babuška and Rheinboldt [3], which involves solving local problems on patches of elements for which the local residuals appear as data.

One issue that frequently complicates residual-based methods is the treatment of boundary conditions on elements or patches in formulating the local problems for error measures. In the equilibrated element residual methods of Ladevèze and Leguillon [11] and Ainsworth and Oden [2], Neumann conditions are used and the residual fluxes are equilibrated for each element. This process can lead to good error estimators, but is generally expensive and rarely used in three-dimensional applications. The use of homogeneous Dirichlet boundary conditions, however, lead to global lower bounds, and often results in estimators that do not deliver acceptable accuracy for use in adaptive mesh strategies.

More recently, variations in the subdomain-residual method have been proposed by Carstensen and Funken [6] and Morin, Nochetto and Siebert [12] (see also Datta [8]), that employ a partition of unity generated by the usual finite element basis

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
2005		**-**

4. TITLE AND SUBTITLE	5a. CONTRACT NUMBER
Analysis of a subdomain-based error estimator for finite element approximations of elliptic problems	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Office of Naval Research,One Liberty Center,875 North Randolph Street Suite 1425,Arlington,VA,22203-1995**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES
The original document contains color images.

14. ABSTRACT
see report

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			
unclassified	**unclassified**	**unclassified**		**33**	

functions for triangular and quadrilateral meshes. By weighting the local bilinear forms using such partitions of unity, local problems for error indicators are formed on the support (patch-domains) of the basis functions, and the weight functions naturally vanish on the patch boundaries. This provides a natural and convenient framework for overcoming problems of assigning boundary averages or equilibrating residual fluxes on the boundaries.

In the present work, we present a detailed analysis of a variant of the method of Morin *et al.* [12] and demonstrate its performance on several model problems on one- and two-dimensional domains. We develop a new computable error estimator and derive both upper and lower bounds on the global error measured in the energy norm. Our numerical experiments suggest the method is easily implemented for elliptic problems and can yield excellent bounds on the global error.

2 Model problem and preliminaries

Let Ω be an open bounded polygonal domain in $\mathbb{R}^d$, $d = 1$ or 2, with boundary $\partial\Omega$ and let u: $\Omega \to \mathbb{R}$ be the solution of the problem

$$\begin{aligned} -\Delta u + cu &= f \quad \text{in } \Omega \\ u &= 0 \quad \text{on } \partial\Omega \end{aligned} \tag{1}$$

where $f \in L^2(\Omega)$ and c is a positive constant. The weak formulation of Problem (1) reads:

$$\text{Find} \quad u \in H_0^1(\Omega) \quad \text{such that} \quad \int_\Omega (\nabla u \cdot \nabla v + cuv) = \int_\Omega fv \quad \forall\ v \in H_0^1(\Omega) \tag{2}$$

where $H_0^1(\Omega)$ is the usual Sobolev space of functions which are in $L^2(\Omega)$ with square integrable derivatives and that vanish on $\partial\Omega$. The bilinear form $B(\cdot,\cdot)$ defined on $H_0^1(\Omega) \times H_0^1(\Omega)$ by

$$B(u,v) = \int_\Omega (\nabla u \cdot \nabla v + cuv) \tag{3}$$

is symmetric, continuous and coercive. It thus defines an inner product in $H_0^1(\Omega)$ and induces the so-called energy norm $|||v||| = \sqrt{B(v,v)}$. Introducing the bounded linear form $F(\cdot)$ defined on $H_0^1(\Omega)$ as

$$F(v) = \int_\Omega fv,$$

the Lax-Milgram Theorem (see e.g. [10]) allows us to establish that the problem:

$$\text{Find} \quad u \in H_0^1(\Omega) \quad \text{such that} \quad B(u,v) = F(v), \quad \forall\ v \in H_0^1(\Omega) \tag{4}$$

is well-posed in the sense that it admits a unique solution $u \in H_0^1(\Omega)$, which depends continuously on the data. Moreover, u is the solution of (1) in the sense of distributions.

Let τ_h be a regular finite element partition of Ω made up of triangular or quadrilateral elements (in 2D). We denote by h_K the diameter of each element K of the mesh τ_h and by ρ_K the diameter of the largest circle inscribed in K, if K is a triangle. A similar definition holds if K is a quadrilateral [1, 7]. The mesh is assumed to be regular, *i.e.* there exists a positive constant $\kappa > 0$ such that:

$$\kappa_K = \frac{h_K}{\rho_K} \leq \kappa \quad \forall\ K \in \tau_h. \tag{5}$$

We then introduce the finite dimensional space $\mathbb{V}_k^h \subset H_0^1(\Omega)$ that is the usual space of $\mathbb{P}_k$ (for meshes of triangles) or $\mathbb{Q}_k$ (for meshes of quadrilaterals) continuous finite elements (see e.g. [7] or [10] for a precise definition). The finite element approximation of problem (4) reads:

$$\text{Find} \quad u_h \in \mathbb{V}_k^h \quad \text{such that} \quad B(u_h, v) = F(v) \quad \forall\ v \in \mathbb{V}_k^h. \tag{6}$$

Denoting the approximation error in u_h by $e = u - u_h$, we have the following equation for the error

$$B(e, v) = B(u - u_h, v) = F(v) - B(u_h, v) \equiv \mathcal{R}(v) \quad \forall\ v \in H_0^1(\Omega). \tag{7}$$

where $\mathcal{R}(.)$ is called the residual functional. Note that the well-known orthogonality property reads:

$$\mathcal{R}(v) = B(e, v) = 0 \quad \forall\ v \in \mathbb{V}_k^h. \tag{8}$$

Lemma 1 *With above assumptions and definitions, the following equality holds:*

$$|||e||| = \sup_{v \in H_0^1(\Omega)} \frac{|\mathcal{R}(v)|}{|||v|||} \equiv \|\mathcal{R}\|_* \tag{9}$$

where $\|\mathcal{R}\|_*$ is the norm of the residual in $H^{-1}(\Omega)$.

Proof: Using (7), we have

$$|||e||| = \frac{B(e, e)}{|||e|||} = \frac{\mathcal{R}(e)}{|||e|||} \leq \sup_{v \in H_0^1(\Omega)} \frac{|\mathcal{R}(v)|}{|||v|||}.$$

Furthermore, from the continuity of $B(\cdot, \cdot)$, $|\mathcal{R}(v)| = |B(e, v)| \leq |||e|||\ |||v|||$ for all $v \in H_0^1(\Omega)$, which establishes the assertion. □

Note that the equality in relation (9) holds because the bilinear form $B(\cdot, \cdot)$ is positive definite and symmetric. In more general cases, we should expect to have only an equivalence between the norm of the residual and the norm of the error. Relation (9) will be useful in deriving a lower bound on the error.

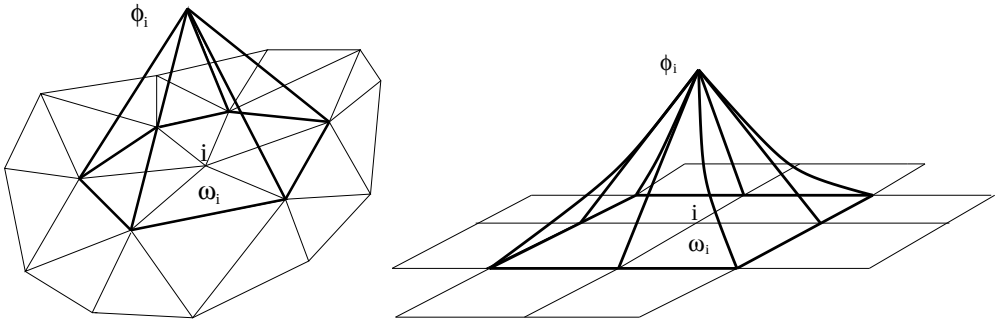

Figure 1: Representation of the function ϕ_i on ω_i for triangular (left) and quadrangular (right) meshes.

3 Upper bound estimate on the error

Let ϕ_i be the set of Lagrangian piecewise linear (for triangles) or piecewise bilinear (for quadrilaterals) basis functions. The support of each ϕ_i is denoted by ω_i and will be referred to as the patch of elements connected to node i of the mesh (see Fig. 1 for a representation of ϕ_i and ω_i in two dimensions). Subsequently, an interior patch ω_i will define a patch for which node i is not located on the boundary $\partial\Omega$. Similarly, a boundary patch ω_i is a patch for which node i is on $\partial\Omega$.

We will use the fact that the functions ϕ_i form a so-called partition of unity

$$\sum_{1\leq i\leq N} \phi_i(x) = 1 \quad \forall\ x \in \Omega \tag{10}$$

where N denotes the total number of nodes of τ_h. Furthermore, let $h_i = \max_{K\in\omega_i} h_K$ and $\rho_i = \min_{K\in\omega_i} \rho_K$. The regularity of the partition τ_h is inherited by the patches ω_i and we have (see Theorem 6.1 in [1]):

$$\exists\ C_0 > 0\colon\ \kappa_i = \frac{h_i}{\rho_i} \leq C_0\ \kappa \quad \forall\ i = 1, ...N. \tag{11}$$

Note that upon inserting (10) into (7), we have

$$B(e,v) = \mathcal{R}(v \sum_{1\leq i\leq N} \phi_i) = \sum_{1\leq i\leq N} \mathcal{R}(v\phi_i) \quad \forall\ v \in H_0^1(\Omega). \tag{12}$$

Let us introduce, now, on each patch ω_i, the bilinear form

$$B_{\phi_i}(u,v) = \int_{\omega_i} \phi_i(\nabla u \cdot \nabla v + cuv) \tag{13}$$

with associated norm $|||v|||_{\phi_i} = \sqrt{B_{\phi_i}(v,v)}$, and define the space for each $i = 1, \ldots, N$ as

$$W(\omega_i) = \overline{\mathcal{W}(\omega_i)}^{\,\|\cdot\|_{\phi_i}}$$

where for an interior patch ω_i:

$$\mathcal{W}(\omega_i) = C^1(\overline{\omega}_i)$$

and for a boundary patch ω_i:

$$\mathcal{W}(\omega_i) = \{v \in C^1(\overline{\omega}_i), \quad v = 0 \ \text{ on } \ \partial\omega_i \cap \partial\Omega\}.$$

Next, we will formulate and solve the following N local problems, over each patch ω_i, for local error functions $\zeta_i \in W(\omega_i)$

$$B_{\phi_i}(\zeta_i, \psi) = \mathcal{R}(\psi\phi_i) \quad \forall\ \psi \in W(\omega_i), \quad 1 \le i \le N. \tag{14}$$

We define the global error estimator

$$\tilde{\varepsilon} = \sqrt{\sum_{1\le i\le N} B_{\phi_i}(\zeta_i, \zeta_i)}. \tag{15}$$

We show in the following that $\tilde{\varepsilon}$ provides for an upper bound on the error $|||e|||$.

Remark 1 *The left-hand side and right-hand side of (14) can be explicitly written as:*

$$\begin{aligned} B_{\phi_i}(\zeta_i, \psi) &= \int_{\omega_i} (\phi_i \nabla\zeta_i \cdot \nabla\psi + c\phi_i\zeta_i\psi), \\ \mathcal{R}(\psi\phi_i) &= \int_{\omega_i} f\psi\phi_i - \int_{\omega_i} (\nabla u_h \cdot \nabla(\psi\phi_i) + cu_h\psi\phi_i). \end{aligned} \tag{16}$$

Whenever ω_i is an interior patch, no boundary conditions are prescribed on $\partial\omega_i$ in (14). Thus, the local problems (14) are of "Neumann" type. In the case $c = 0$, the solution is defined up to a constant and we should expect to have a compatibility condition on the right-hand side. Indeed, thanks to the Galerkin orthogonality (8), it happens that for ψ a constant function on ω_i

$$\mathcal{R}(\psi\phi_i) = \psi\mathcal{R}(\phi_i) = 0. \tag{17}$$

Thus, the compatibility condition is always satisfied. In order to fix the constant, we seek a solution of (14) in the space

$$\overset{\circ}{W}(\omega_i) = \{v \in W(\omega_i), \quad \int_{\omega_i} v\phi_i = 0\} \tag{18}$$

instead of $W(\omega_i)$.

Remark 2 *As written in (16), the residual does not contain any integrals of the fluxes along the element interfaces. As presented, it is shown that it is not necessary to perform the integration by parts as most authors do [6, 12].*

In Lemma 3, we will prove the well-posedness of the local problems (14). Toward this goal, we state Lemma 2 below. Its proof has been given in [12] for a triangular mesh and is provided in the Appendix in the case of quadrangular meshes. We remark that the proof varies substantially when dealing with triangles or quadrilaterals. Indeed, the weighting function ϕ_i is linear for the first type of elements. For quadrilateral elements, ϕ_i is bilinear which means that the function can be quadratic along certain directions in the elements, making the proof a little more complex.

Lemma 2 (Weighted Poincaré inequality) *Let h_i be the maximum size of the elements in the patch ω_i. For any function $\zeta_i \in W(\omega_i)$ that satisfies the condition $\int_{\omega_i} \zeta_i \phi_i = 0$ whenever ω_i is an interior patch, there exists a constant C, independent of h_i, but dependent of κ, such that*

$$\|\zeta_i\|_{L^2(\omega_i)} \leq C h_i \|\nabla \zeta_i\|_{\phi_i}, \tag{19}$$

where $\|\cdot\|_{\phi_i} = \|\cdot \phi_i^{1/2}\|_{L^2(\omega_i)}$.

Lemma 3 *Problem (14) is well posed for any given $h_i > 0$.*

Proof: We first proof this lemma for an interior patch ω_i. The bilinear form $B_{\phi_i}(\cdot,\cdot)$ is an inner product on $W(\omega_i)$ and thus is continuous and coercive. The assertion then follows from the Riesz representation theorem if we can show that the functional $\mathcal{R}(\psi\phi_i)$ is bounded in $W(\omega_i)$ with respect to the norm $|||\psi|||_{\phi_i}$.

Let us first remark that if $c = 0$, all functions $\psi \in \overset{\circ}{W}(\omega_i)$ satisfy the condition $\int_{\omega_i} \psi\phi_i = 0$. On the other hand, if $c \neq 0$, thanks to the Galerkin orthogonality (8), we have $\mathcal{R}(\phi_i) = 0$ and, for all $\psi \in W(\omega_i)$,

$$\mathcal{R}(\psi\phi_i) = \mathcal{R}\left((\psi - \frac{\int_{\omega_i}\psi\phi_i}{\int_{\omega_i}\phi_i})\phi_i\right) = \mathcal{R}(\tilde{\psi}\phi_i),$$

where $\tilde{\psi}$ satisfies now the condition $\int_{\omega_i}\tilde{\psi}\phi_i = 0$. Hence,

$$\begin{aligned}
|\mathcal{R}(\psi\phi_i)| = |\mathcal{R}(\tilde{\psi}\phi_i)| &= \left|\int_{\omega_i} f\tilde{\psi}\phi_i - \int_{\omega_i}(\nabla u_h \cdot \nabla(\tilde{\psi}\phi_i) + c u_h \tilde{\psi}\phi_i)\right| \\
&\leq \left|\int_{\omega_i} f\tilde{\psi}\phi_i\right| + \left|\int_{\omega_i}(\phi_i \nabla u_h \cdot \nabla\tilde{\psi} + c u_h\tilde{\psi}\phi_i)\right| + \left|\int_{\omega_i}\tilde{\psi}\nabla u_h \cdot \nabla\phi_i\right| \\
&\leq \|f\|_{\phi_i}\|\tilde{\psi}\|_{\phi_i} + \|\nabla u_h\|_{\phi_i}\|\nabla\tilde{\psi}\|_{\phi_i} + c\|u_h\|_{\phi_i}\|\tilde{\psi}\|_{\phi_i} + \left|\int_{\omega_i}\tilde{\psi}\nabla u_h\cdot\nabla\phi_i\right|.
\end{aligned} \tag{20}$$

Let us remark now that

$$\|\tilde{\psi}\|_{\phi_i} \leq 2\|\psi\|_{\phi_i}. \tag{21}$$

Indeed, if we set $g = \int_{\omega_i} \psi\phi_i / \int_{\omega_i} \phi_i$, we have

$$\|g\|_{\phi_i}^2 = g^2 \int_{\omega_i} \phi_i = \frac{(\int_{\omega_i} \psi\phi_i)^2}{\int_{\omega_i} \phi_i} \leq \frac{\|\psi\|_{\phi_i}^2 (\int_{\omega_i} \phi_i)}{\int_{\omega_i} \phi_i} = \|\psi\|_{\phi_i}^2$$

and clearly,

$$\|\tilde{\psi}\|_{\phi_i} \leq \|\psi\|_{\phi_i} + \|g\|_{\phi_i} \leq 2\|\psi\|_{\phi_i}.$$

Thus, noticing that $\nabla\tilde{\psi} = \nabla\psi$ and using the previous inequality, we have

$$|\mathcal{R}(\psi\phi_i)| \leq 2\|f\|_{\phi_i}\|\psi\|_{\phi_i} + \|\nabla u_h\|_{\phi_i}\|\nabla\psi\|_{\phi_i} + 2c\|u_h\|_{\phi_i}\|\psi\|_{\phi_i} + \left|\int_{\omega_i} \tilde{\psi}\nabla u_h \cdot \nabla\phi_i\right|.$$

We now need to bound the term $\left|\int_{\omega_i} \tilde{\psi}\nabla u_h \cdot \nabla\phi_i\right|$. We have by Hölder's inequality:

$$\left|\int_{\omega_i} \tilde{\psi}\nabla u_h \cdot \nabla\phi_i\right| \leq \|\nabla\phi_i\|_{L^\infty(\omega_i)} \, \|\nabla u_h\|_{L^2(\omega_i)} \, \|\tilde{\psi}\|_{L^2(\omega_i)}$$

Moreover, a well-known property of Lagrangian basis functions allows us to state that

$$\|\nabla\phi_i\|_{L^\infty(K)} \leq C_1 \frac{\kappa_K}{\rho_K} \quad \forall K \in \omega_i \quad \text{for } K \text{ quadrilaterals,}$$
$$\|\nabla\phi_i\|_{L^\infty(K)} \leq \frac{C_2}{\rho_K} \quad \forall K \in \omega_i \quad \text{for } K \text{ triangles.}$$

In both cases,

$$\|\nabla\phi_i\|_{L^\infty(\omega_i)} = \max_{K\in\omega_i} \|\nabla\phi_i\|_{L^\infty(K)} \leq \frac{C'}{\rho_i} \tag{22}$$

where $\rho_i = \min_{K\in\omega_i}(\rho_K)$ and C' might depend on the regularity parameter κ of the partition τ_h.

Thus,

$$\begin{aligned}\left|\int_{\omega_i} \tilde{\psi}\nabla u_h \nabla\phi_i\right| &\leq \frac{C'}{\rho_i}\|\nabla u_h\|_{L^2(\omega_i)} \, \|\tilde{\psi}\|_{L^2(\omega_i)} \\ &\leq \frac{C'Ch_i}{\rho_i}\|\nabla u_h\|_{L^2(\omega_i)} \, \|\nabla\tilde{\psi}\|_{\phi_i} \\ &\leq \frac{C'Ch_i}{\rho_i}\|\nabla u_h\|_{L^2(\omega_i)} \, \|\nabla\psi\|_{\phi_i}\end{aligned}$$

where we have used the weighted Poincaré inequality (19).

Then, we have

$$\begin{aligned}|\mathcal{R}(\psi\phi_i)| \leq{}& 2\|f\|_{\phi_i}\|\psi\|_{\phi_i} + \|\nabla u_h\|_{\phi_i}\|\nabla\psi\|_{\phi_i} + 2c\|u_h\|_{\phi_i}\|\psi\|_{\phi_i} \\ &+ \frac{C'Ch_i}{\rho_i}\|\nabla u_h\|_{L^2(\omega_i)} \, \|\nabla\psi\|_{\phi_i}\end{aligned}$$

and noticing that $\|\nabla\psi\|_{\phi_i} \leq |||\psi|||_{\phi_i}$ and $\|\psi\|_{\phi_i} \leq c^{-1/2}|||\psi|||_{\phi_i}$, we finally obtain

$$|\mathcal{R}(\psi\phi_i)| \leq \left(\frac{2}{\sqrt{c}}\|f\|_{\phi_i} + \|\nabla u_h\|_{\phi_i} + 2\sqrt{c}\,\|u_h\|_{\phi_i} + \frac{C'Ch_i}{\rho_i}\|\nabla u_h\|_{L^2(\omega_i)}\right)|||\psi|||_{\phi_i} \tag{23}$$

The proof is thus complete since all the quantities appearing on the right-hand side are bounded. The same proof is also valid for a boundary patch, without having to introduce the function $\tilde{\psi}$. □

We now derive an exact upper bound for the error. Using the previous definitions and the Cauchy-Schwarz inequality, we have for all $v \in H_0^1(\Omega)$

$$\begin{aligned} B(e,v) &= \sum_{1\leq i\leq N} \mathcal{R}(v\phi_i) \\ &= \sum_{1\leq i\leq N} B_{\phi_i}(\zeta_i, v) \\ &\leq \sum_{1\leq i\leq N} \sqrt{B_{\phi_i}(\zeta_i,\zeta_i)}\sqrt{B_{\phi_i}(v,v)} \\ &\leq \sqrt{\sum_{1\leq i\leq N} B_{\phi_i}(\zeta_i,\zeta_i)}\sqrt{\sum_{1\leq i\leq N} B_{\phi_i}(v,v)}. \end{aligned}$$

Taking $v = e$ in the previous inequality, noticing that $\sum_{1\leq i\leq N} B_{\phi_i}(v,v) = B(v,v)$ for any v in $H_0^1(\Omega)$, and using the definition of $\tilde{\varepsilon}$ in (15), we obtain the following guaranteed upper bound on the error, *i.e.*

$$|||e||| = \sqrt{B(e,e)} \leq \tilde{\varepsilon}. \tag{24}$$

However, the quantity $\tilde{\varepsilon}$ cannot be computed exactly because the local problems (14) defined on each patch for ζ_i are of infinite dimension. We show in the next section how to approximate these problems in order to derive an *a posteriori* error estimator.

4 The *a posteriori* error estimator

Let $P^{k+p}(\omega_i)$ denote the finite element space of piecewise polynomials of degree $k+p$ on the patch ω_i and let $\mathcal{P}^{k+p}(\omega_i) = P^{k+p}(\omega_i) \cap W(\omega_i)$. Here, p represents the extra degree with respect to the degree k used to compute the finite element solution u_h. On each patch ω_i, we now solve for $\eta_i \in \mathcal{P}^{k+p}(\omega_i)$ as the solution to the problem

$$B_{\phi_i}(\eta_i, \psi) = \mathcal{R}(\psi\phi_i) \quad \forall\ \psi \in \mathcal{P}^{k+p}(\omega_i) \tag{25}$$

and define the global error estimator

$$\varepsilon = \sqrt{\sum_{1\le i\le N} \varepsilon_i^2} = \sqrt{\sum_{1\le i\le N} B_{\phi_i}(\eta_i,\eta_i)}. \tag{26}$$

The quantities $\varepsilon_i = \sqrt{B_{\phi_i}(\eta_i,\eta_i)}$ represent the local contribution from each patch ω_i to the global error estimator ε.

Remark 3 *Although the residual vanishes on $\mathbb{V}_k^h$, we can actually choose $p = 0$ in (25) because the term $\mathcal{R}(\psi\phi_i)$ does not necessarily vanish for $\psi \in \mathcal{P}^k(\omega_i)$ due to the presence of the weighting function ϕ_i. However, we expect to get more accurate estimates as p is increased. Indeed we can verify from Problem (25) that the error estimator ε should increase as the value of p increases since:*

$$\varepsilon_i = |||\eta_i|||_{\phi_i} = \sup_{v\in\mathcal{P}^{k+p}(\omega_i)} \frac{|\mathcal{R}(v\phi_i)|}{|||v|||_{\phi_i}}. \tag{27}$$

Lemma 4 *Let $\mathbb{V}_h^{k+p}(\Omega)$ be the global finite element space of piecewise polynomials of degree $k+p$ defined on Ω and vanishing on $\partial\Omega$. Then we have*

$$|||e||| \le \varepsilon + 2 \inf_{v\in\mathbb{V}_h^{k+p}(\Omega)} |||u-v|||. \tag{28}$$

Proof: On one hand, we have for all $v \in \mathbb{V}_h^{k+p}(\Omega)$,

$$\begin{aligned} |||v-u_h|||^2 &= B(v-u_h, v-u_h) \\ &= B(e, v-u_h) + B(v-u, v-u_h) \\ &\le B(e, v-u_h) + |||v-u||| \; |||v-u_h|||. \end{aligned} \tag{29}$$

On the other hand, using (7) and the Cauchy-Schwarz inequality, we have

$$\begin{aligned} B(e, v-u_h) &= \mathcal{R}(v-u_h) \\ &= \sum_{1\le i\le N} \mathcal{R}((v-u_h)\phi_i) \\ &= \sum_{1\le i\le N} \mathcal{R}((v-u_h)_{|\omega_i}\phi_i) \end{aligned} \tag{30}$$

where $(v-u_h)_{|\omega_i}$ denotes the restriction of $(v-u_h)$ on ω_i. Since $(v-u_h)_{|\omega_i} \in \mathcal{P}^{k+p}(\omega_i)$ and using (25), we may further write

$$\begin{aligned} B(e, v-u_h) &= \sum_{1\le i\le N} B_{\phi_i}(\eta_i, (v-u_h)_{|\omega_i}) \\ &\le \sqrt{\sum_{1\le i\le N} B_{\phi_i}(\eta_i,\eta_i)} \sqrt{\sum_{1\le i\le N} B_{\phi_i}(v-u_h, v-u_h)} \\ &\le \varepsilon |||v-u_h|||. \end{aligned} \tag{31}$$

With (29) and (31), we obtain $|||v-u_h||| \leq \varepsilon + |||v-u|||$ and $|||e||| \leq |||u-v||| + |||v-u_h||| \leq \varepsilon + 2|||u-v|||$. Since the last inequality holds for all $v \in \mathbb{V}_h^{k+p}(\Omega)$, inequality (28) is thus established. □

Inequality (28) shows that the estimator ε does not necessarily provide for an upper bound on the error. However, the quantity $2 \inf_{v \in \mathbb{V}_h^{k+p}(\Omega)} |||u-v|||$ should decrease as p increases, so that ε should be an upper bound for high enough p.

As a final result, we show that the estimator ε is equivalent to the energy norm of the error.

Lemma 5 *There exists a constant $\tilde{C} > 1$, that depends only on the regularity of the mesh and on the constant C appearing in Lemma 2, but otherwise independent of h and p, such that*

$$\varepsilon \leq \tilde{C}|||e|||. \tag{32}$$

Proof: For any $i = 1, \ldots, N$, the function $\eta_i \phi_i$ is a piecewise polynomial function that vanishes on the boundary $\partial\omega_i$. Let us denote by $\widetilde{\eta_i\phi_i}$ the extension of $\eta_i\phi_i$ by zero over $\Omega \setminus \omega_i$. Clearly, $\widetilde{\eta_i\phi_i}$ belongs to $H_0^1(\Omega)$ and can be taken as a test function in the continuous problem (4). Moreover, its gradient will also vanish outside ω_i. We thus have,

$$\int_{\omega_i} f\eta_i\phi_i = \int_\Omega f\widetilde{\eta_i\phi_i} = B(u, \widetilde{\eta_i\phi_i}) = \int_{\omega_i} \left(\nabla u \cdot \nabla(\eta_i\phi_i) + cu\eta_i\phi_i\right).$$

Hence,

$$\begin{aligned} |||\eta_i|||_{\phi_i}^2 = \mathcal{R}(\eta_i\phi_i) &= \int_{\omega_i} f\eta_i\phi_i - \int_{\omega_i} \left(\nabla u_h \cdot \nabla(\eta_i\phi_i) + cu_h\eta_i\phi_i\right) \\ &= \int_{\omega_i} \left[\nabla(u-u_h)\cdot\nabla(\eta_i\phi_i) + c(u-u_h)\eta_i\phi_i\right] \\ &= B_{\phi_i}(u-u_h, \eta_i) + \int_{\omega_i} \nabla(u-u_h)\cdot\nabla\phi_i\eta_i. \end{aligned} \tag{33}$$

Let us remark now that the local solution η_i belongs to the space $\overset{\circ}{W}(\omega_i)$ also in the case $c > 0$. Indeed, by taking a constant test function $\psi = \bar{c}$ in (25) we have

$$B_{\phi_i}(\eta_i, \bar{c}) = c\int_{\omega_i} \eta_i\bar{c}\phi_i = \mathcal{R}(\bar{c}\phi_i) = \bar{c}\ \mathcal{R}(\phi_i) = 0,$$

owing to the Galerkin orthogonality. The local functions η_i then satisfy the weighted Poincaré inequality,

$$\|\eta_i\|_{L^2(\omega_i)} \leq Ch_i\|\nabla\eta_i\|_{\phi_i}.$$

We can proceed in a similar way as in the proof of Lemma 3 to bound the last term in (33), thus obtaining

$$|||\eta_i|||_{\phi_i}^2 \leq |||u-u_h|||_{\phi_i}|||\eta_i|||_{\phi_i} + C\|\nabla(u-u_h)\|_{L^2(\omega_i)}\|\nabla\eta_i\|_{\phi_i},$$

where the constant C depends on the constant appearing in the weighted Poincaré inequality and on the regularity parameter κ of the mesh. Summing over all the patches ω_i and using the discrete Cauchy-Schwarz inequality, we have

$$\varepsilon^2 \leq \sqrt{\sum_{1\leq i\leq N} |||u-u_h|||^2_{\phi_i}} \sqrt{\sum_{1\leq i\leq N} \varepsilon_i^2} + C \sqrt{\sum_{1\leq i\leq N} \|\nabla(u-u_h)\|^2_{L^2(\omega_i)}} \sqrt{\sum_{1\leq i\leq N} \varepsilon_i^2}.$$

That is:

$$\varepsilon \leq |||u-u_h||| + C \sqrt{\sum_{1\leq i\leq N} \|\nabla(u-u_h)\|^2_{L^2(\omega_i)}}. \tag{34}$$

For each element K of the mesh τ_h, let us denote $N_K = \{i : \ K \subset \omega_i\}$, $\lambda_K = \dim N_K$ and $\lambda = \max_{K\in\tau_h} \lambda_K$. Then, we have

$$\begin{aligned} \sum_{1\leq i\leq N} \|\nabla(u-u_h)\|^2_{L^2(\omega_i)} &= \sum_{1\leq i\leq N} \sum_{K\subset\omega_i} \|\nabla(u-u_h)\|^2_{L^2(K)} \\ &= \sum_{K\subset\tau_h} \lambda_K \|\nabla(u-u_h)\|^2_{L^2(K)} \leq \lambda |||u-u_h|||^2 \end{aligned} \tag{35}$$

and the final result (32) is achieved with $\tilde{C} = 1 + C\sqrt{\lambda}$. □

5 Recovery of a global lower bound on the error

Relation (9) shows that for any function $z \in H^1_0(\Omega)$, the quantity $|\mathcal{R}(z)|/|||z|||$ provides a lower bound of the error. Here we choose to construct the function z as

$$z = \sum_{1\leq i\leq N} \eta_i \phi_i,$$

η_i being the solution of problems (25). The function z is continuous since the basis functions ϕ_i vanish at the boundaries of all patches; hence, z is in $H^1_0(\Omega)$. This choice is motivated by the fact that the functions η_i are already available upon solving the local problems (25).

Furthermore, we note from (25) that:

$$\mathcal{R}(z) = \mathcal{R}(\sum_{1\leq i\leq N} \eta_i\phi_i) = \sum_{1\leq i\leq N} \mathcal{R}(\eta_i\phi_i) = \sum_{1\leq i\leq N} B(\eta_i,\eta_i) = \varepsilon^2.$$

The lower bound on the error then reduces to:

$$\eta_{low} = \frac{\varepsilon^2}{|||z|||}. \tag{36}$$

In other words, since ε^2 is already known, it suffices to compute the norm $|||z||| = |||\sum_{1\leq i\leq N} \eta_i\phi_i|||$ to derive a lower bound on the error.

Remark 4 *As observed in Remark 1, in the case $c = 0$, the solutions η_i of problem (25) on an interior patch ω_i are obviously defined up to an arbitrary constant. We note that the upper bound*

$$\varepsilon^2 = \sum_{1 \leq i \leq N} \int_{\omega_i} |\nabla \eta_i|^2 \phi_i$$

does not depend on this "arbitrary" constant. On the other hand, for the lower bound η_{low}, the quantity

$$|||z||| = \sqrt{\int_\Omega |\nabla (\sum_{1 \leq i \leq N} \eta_i \phi_i)|^2} = \sqrt{\int_\Omega (\sum_{1 \leq i \leq N} (\nabla \eta_i \phi_i + \eta_i \nabla \phi_i))^2}$$

does depend on the constant. Following the suggestion in Remark 1, we choose the constant c_i as:

$$c_i = \frac{\int_{\omega_i} \eta_i \phi_i}{\int_{\omega_i} \phi_i}. \tag{37}$$

In this manner, $\eta_i - c_i \in \mathcal{P}^{k+p}(\omega_i) \cap \overset{\circ}{W}(\omega_i)$ and the function z is defined as:

$$z = \sum_{1 \leq i \leq N} (\eta_i - c_i) \phi_i. \tag{38}$$

6 A strategy for mesh adaptation

In addition to accuracy assessment, *a posteriori* error estimators can be used as refinement indicators for mesh adaptation in order to reduce the numerical error. A simple strategy consists of defining local contributions to the global error and refining the elements which exhibit large source of error.

For the subdomain-based error estimator proposed here, there are actually two ways for splitting the global error into sums of local contributions: namely, we can define patch-wise or element-wise contributions. From (26), we have

$$\varepsilon = \sqrt{\sum_{1 \leq i \leq N} \varepsilon_i^2}$$

where $\varepsilon_i = \sqrt{B_{\phi_i}(\eta_i, \eta_i)}$ represents the contribution of the patch ω_i to the error estimate ε. Similarly, we can decompose ε into a sum of element-wise contributions.

Indeed,

$$\begin{aligned}\varepsilon = \sqrt{\sum_{1\le i\le N} \varepsilon_i^2} &= \sqrt{\sum_{1\le i\le N} \int_{\omega_i} (|\nabla\eta_i|^2 + c\eta_i^2)\phi_i} \\ &= \sqrt{\sum_{1\le i\le N} \left(\sum_{K\subset\omega_i} \int_K (|\nabla\eta_i|^2 + c\eta_i^2)\phi_i \right)} \\ &= \sqrt{\sum_{K\subset\tau_h} \left(\sum_{i\in N_K} \int_K (|\nabla\eta_i|^2 + c\eta_i^2)\phi_i \right)} \\ &= \sqrt{\sum_{K\subset\tau_h} \varepsilon_K^2}\end{aligned}$$

where, now, the quantity ε_K represents the element-wise contribution. The procedure can naturally be applied to the lower bound estimator as well; in fact, we have:

$$\eta_{low} = \frac{\varepsilon^2}{|||z|||} = \sum_{1\le K\le N_e} \frac{\varepsilon_K^2}{|||z|||} = \sum_{1\le K\le N_e} \varepsilon_{K,low}$$

with $\varepsilon_{K,low} = \varepsilon_K^2/|||z|||$ the local contribution per element.

In the numerical experiments shown in the next sections, we only consider refinements based on the element-wise contributions (in order to be able to compare our mesh adaptation procedure with the element residual method). Namely, an element $K \subset \tau_h$ is refined if either

$$\frac{\varepsilon_K}{\max\limits_{K\subset\tau_h}(\varepsilon_K)} \ge C_{adap} \qquad \text{or} \qquad \frac{\varepsilon_{K,low}}{\max\limits_{K\subset\tau_h}(\varepsilon_{K,low})} \ge C_{adap} \tag{39}$$

where C_{adap} is a user-prescribed constant parameter ranging from zero to one. In the numerical tests presented in section 8, we have chosen $C_{adap} = 0.5$.

7 Numerical experiments in 1D

The primary objective in these examples is to assess the performance of the upper and lower bound estimates on the numerical error. In all experiments, the quality of these estimates will be measured in terms of the effectivity index, namely the ratio between the error estimate and the exact error in the energy norm $|||\cdot|||$.

We first consider the two–point boundary–value problem:

$$\begin{aligned} -u'' + cu &= f \quad \text{in } (0,1) \\ u &= 0 \quad \text{at } x = 0 \\ u' &= 0 \quad \text{at } x = 1 \end{aligned} \tag{40}$$

where the constant c will be chosen as either one or zero.

7.1 Problem with smooth solution

In our first example, we select the load f such that the exact solution (see Fig. 2) is analytically given by:

$$u(x) = \frac{27}{4}\left((1-x)(1-3x) - 2x(x-\frac{1}{3})(1-x)^2\right) e^{-(x-1/3)^2}.$$

The finite element approximations are computed using polynomial degrees k =1, 2 or 3 and, for each case, we calculate the upper bound ε and lower bound η_{low} setting the extra degree as $p = 0,\ 1,\ 2$ or 3. The effectivity indices are shown in Figs. 3, 4 and 5 respectively on sequences of uniformly refined meshes.

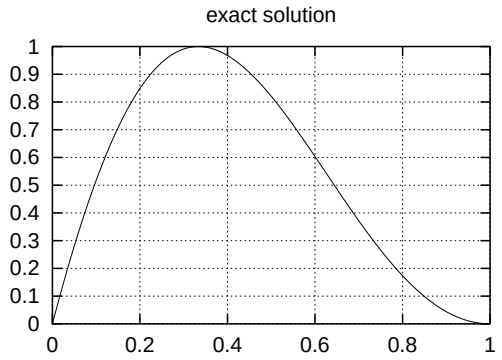

Figure 2: Representation of the exact solution $u(x)$.

We remark first that the effectivity indices are nearly identical for $p = 1$, 2 or 3. In other words, this shows that the local problems rapidly "saturate" as p is increased. More surprisingly, we also observe that most of the results are acceptable when taking $p = 0$. This is an important observation as in many existing finite element codes, there is no access to polynomial shape functions of higher degree than the ones used for the finite element solution u_h. This aspect of the error estimator may be an advantage over other element residual type methods for which it is necessary to take $p > 0$. However, if we choose $p = 0$, we are not guaranteed that the quantity ε provides for an upper bound on the error (even as h tends to zero). In this case, more sophisticated techniques may be envisaged, that aim at recovering a guaranteed upper bound by estimating the error $\|\zeta_i - \eta_i\|_{\phi_i}$ in the solution of the local problems, as well, as proposed in [4, 13]. We will not investigate this approach in the present work.

Another observation deals with the fact that the error bounds barely depend on the mesh size h (at least when the mesh is not coarse). Furthermore, the effectivity index of the upper bound decreases as k goes from one to three. In fact, we find an effectivity index close to 1.22 for $k = 1$, 1.07 for $k = 2$ and 1.04 for $k = 3$. On the other hand, the lower bound η_{low} always provides for effectivity indices very close to one (even for $p = 0$) even for coarse meshes..

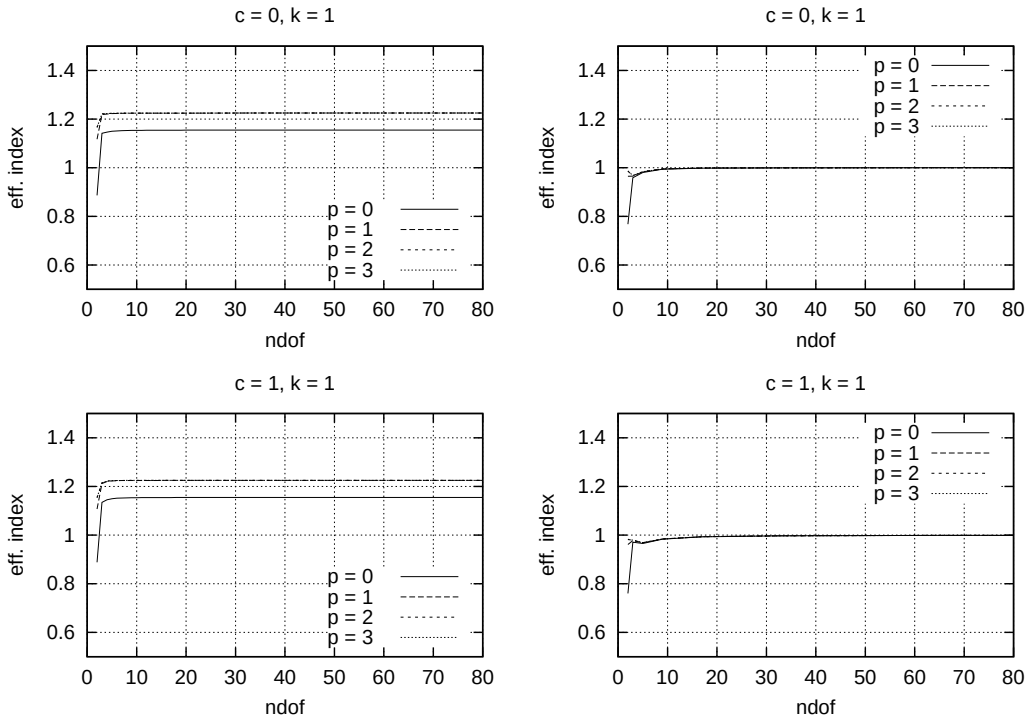

Figure 3: Global effectivity index for the upper (left) and lower (right) bounds for the case $k = 1$ and $c = 0$ (top), $c = 1$ (bottom).

7.2 A variant of the error estimator

In [12], the authors proposed to solve the local problems (25) using homogeneous Dirichlet boundary conditions. In other words, defining $\mathcal{P}_0^{k+p}(\omega_i)$ as the subspace of $\mathcal{P}^{k+p}(\omega_i)$ such as:

$$\mathcal{P}_0^{k+p}(\omega_i) = \{v \in \mathcal{P}^{k+p}(\omega_i), v = 0 \text{ on } \partial\omega_i\},$$

they solve for $\eta_i \in \mathcal{P}_0^{k+p}(\omega_i)$ such that

$$B_{\phi_i}(\eta_i, v) = \mathcal{R}(v\phi_i) \quad \forall\ v \in \mathcal{P}_0^{k+p}(\omega_i). \tag{41}$$

Here we compare the error estimators (upper and lower bounds) when using either Dirichlet or Neumann boundary conditions. For the case $c = 1$, $k = 1$, and taking $p = 0$ or 1, we obtain the results shown in Fig. 6.

We observe that the effectivity indices for both the upper and lower bounds converge at a slower rate with respect to the number of degrees of freedom when using homogeneous Dirichlet boundary conditions. We conclude that the subdomain-based error estimation method is more reliable when we do not impose any boundary conditions on the patch boundaries.

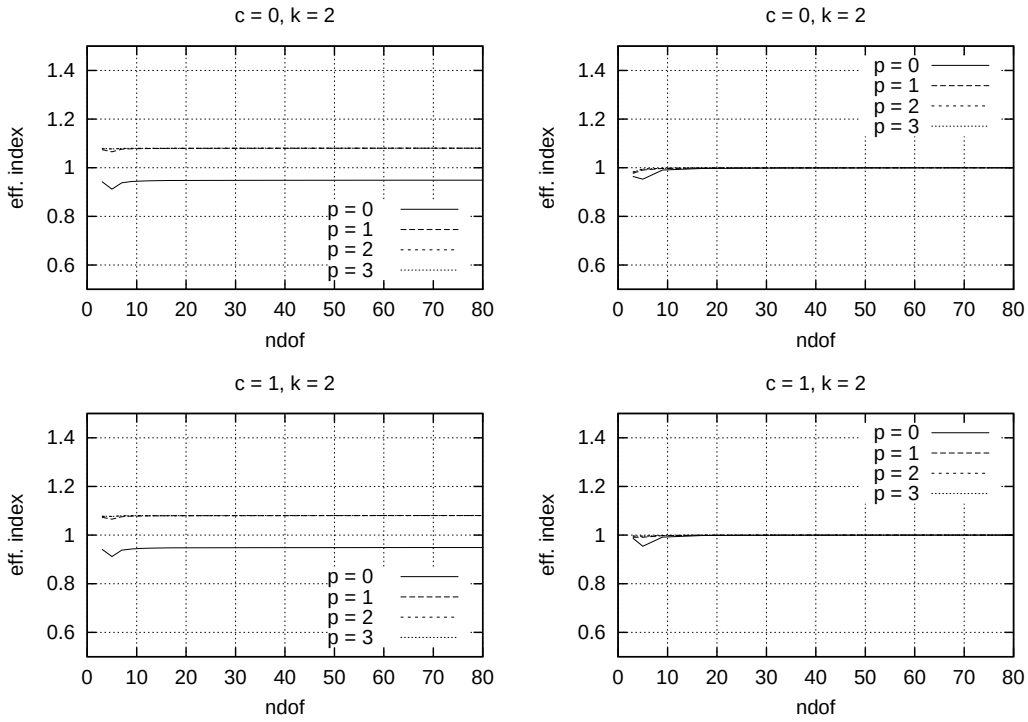

Figure 4: Global effectivity index for the upper (left) and lower (right) bounds for the case $k = 2$ and $c = 0$ (top), $c = 1$ (bottom).

7.3 Comparison with other error estimators

In this section, we compare the subdomain-based error estimator (SRM for Subdomain Residual Method) with other methods: namely, the ZZ recovery-type method [14, 15], and the Element Residual Method (ERM) [5, 9]. We choose here to solve (40) with $c = 0$. Note that the results are in this case the same whether we use either ERM or the Element Residual Method with the equilibration technique (EqRM) [11, 1].

We want to test these three estimators using an approximation of the Green function as a solution of (40). We recall that the Green function is the solution of (40) (with $c = 0$) when the load f is the Dirac distribution. The Green function is particularly interesting because it possesses a discontinuity in the first derivative at the point where the Dirac distribution is located. To compute a solution of (40) which approximates the Green function, we actually replace the Dirac distribution located at $x = x_0$ by the function $f = k_\epsilon$ defined by:

$$k_\epsilon(x) = \begin{cases} C \exp\left(\dfrac{|x - x_0|^2}{\epsilon^2} - 1\right)^{-1} & \text{if } |x - x_0| < \epsilon, \\ 0 & \text{if } |x - x_0| \geq \epsilon \end{cases} \tag{42}$$

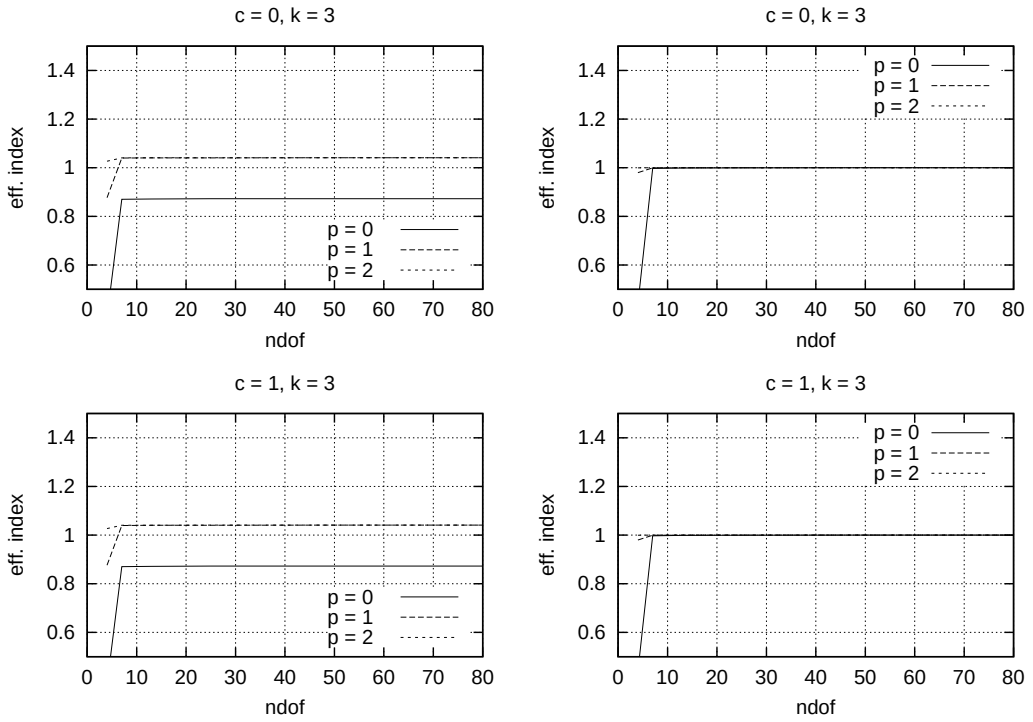

Figure 5: Global effectivity index for the upper (left) and lower (right) bounds for the case $k = 3$ and $c = 0$ (top), $c = 1$ (bottom).

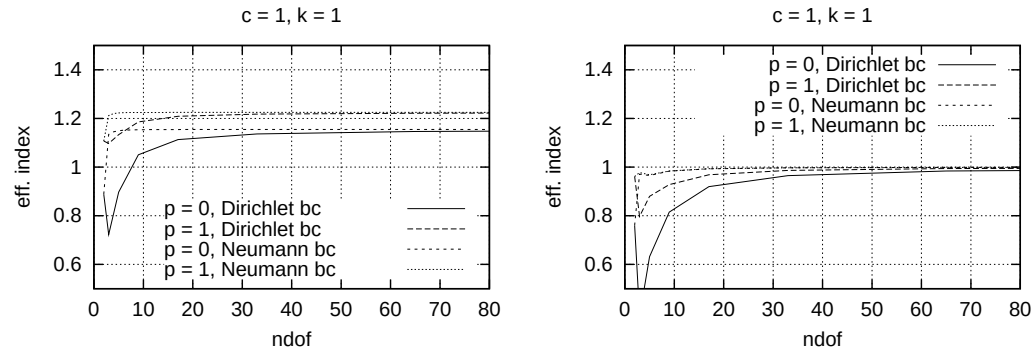

Figure 6: Global effectivity index for the upper and lower bounds, Dirichlet or Neumann boundary conditions.

where the constant C is chosen to satisfy

$$\int_\Omega k_\epsilon(x)\,dx = 1. \tag{43}$$

In the following, we take $\epsilon = 0.1$ for which the corresponding load f and "exact"

solution u are plotted in Fig. 7. In this case, the exact solution is not known and is approximated by an overkilled finite element solution $\tilde{u}_h$ using 1500 degrees of freedom.

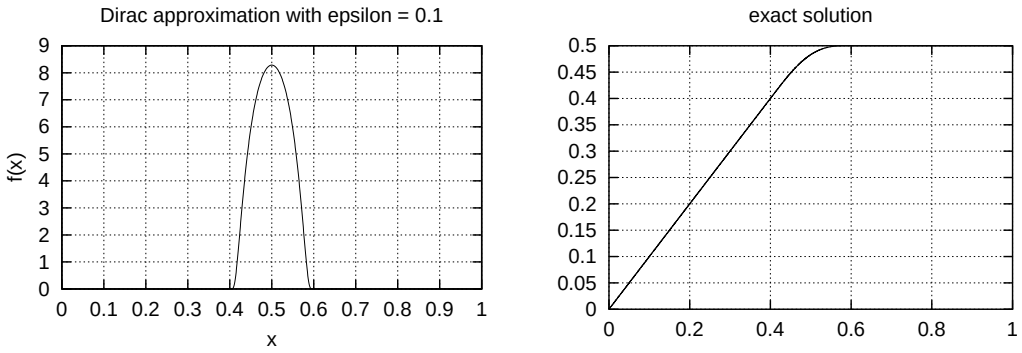

Figure 7: Representation of the load (left) and the exact solution (right).

In order to compute effectivity indices, we also calculate an accurate value of the energy norm of u such as:

$$|||u||| \approx \sqrt{|||\tilde{u}_h|||^2 + |||\tilde{e}|||^2}$$

where $\tilde{e}$ is the error in $\tilde{u}_h$ estimated here by ERM. The "exact" error e in the finite element solution u_h will then be calculated by the formula (obtained from the orthogonality property (8)):

$$|||e|||^2 = |||u|||^2 - |||u_h|||^2 \approx |||\tilde{u}_h|||^2 + |||\tilde{e}|||^2 - |||u_h|||^2 \tag{44}$$

In Fig. (8), we present the different results obtained with the three error estimators. In each case, the exact error is estimated using (44). We take $k = 1$ for the degree of the finite element solution, $p = 1$ for the extra-degree used in SRM and ERM. The first three graphs represent the distribution of the error estimate on the domain $[0, 1]$ for SRM, ERM, and ZZ using 30 elements. We observe that the approximation error is concentrated in the middle elements, where the main variation in the first derivative of the exact solution actually occurs. The last graph represents the behavior of the effectivity index for these three estimators when the number of degrees of freedom is increased.

Our main conclusion is that the ZZ method is poorly accurate when we do not use enough degrees of freedom (less than 50 degrees of freedom in our example). This is due to the fact that the ZZ method is based on a recovery of the gradient of the solution u_h. However, the two other estimators seem to be quite stable. The effectivity index with SRM stays close to 1.22, which is the value we obtained in our previous experiments, whereas the effectivity index using ERM remains close to unity.

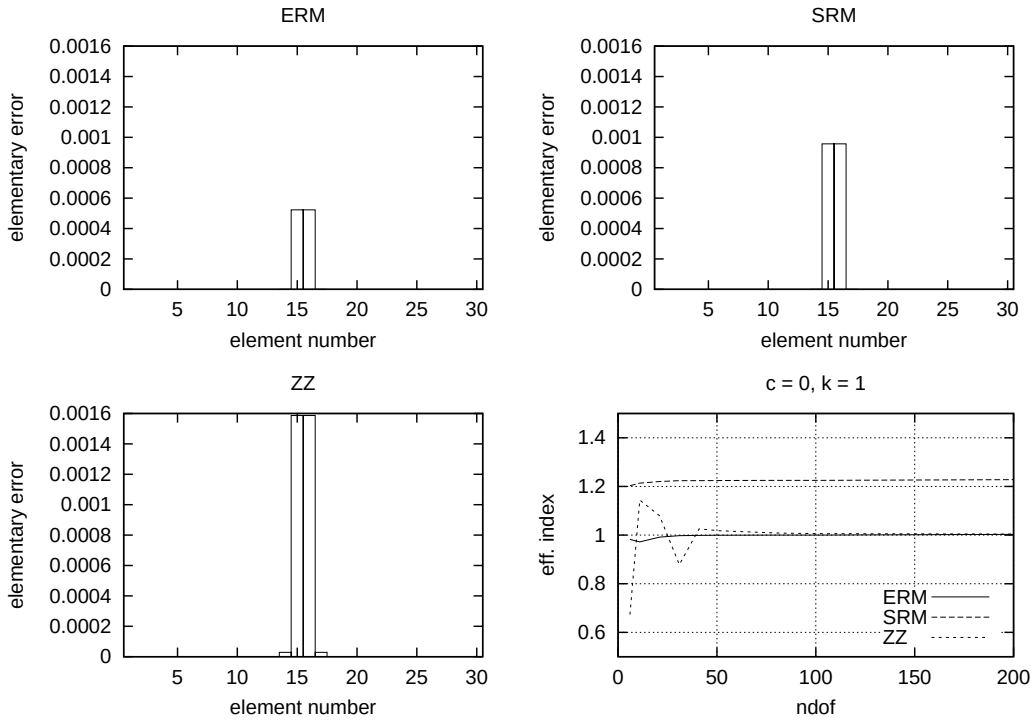

Figure 8: Distribution of the error estimate with ERM (top left), SRM (top right) and the ZZ (bottom left). Effectivity indices for each method (bottom right).

8 Numerical experiments in 2D

8.1 Performance of the estimator

We now investigate the performance of the estimator in the case of 2D problems. We consider the squared domain $\Omega = \{(x, y) \in \mathbb{R}^2,\ 0 < x < 1,\ 0 < y < 1\}$ and solve the problem

$$\begin{aligned} -\Delta u + cu &= f \quad \text{in } \Omega \\ u &= 0 \quad \text{on } \partial\Omega \end{aligned} \tag{45}$$

where c is a constant chosen as either 0 or 1. The load f is such that the exact solution, plotted in Fig. 9, is given by:

$$u(x) = 0.0005\ x^2(1-x)^2y^2(1-y)^2e^{10(x^2+y)}. \tag{46}$$

The following figures present the results for different values of the parameters: the degree of the finite element solution is chosen as $k = 1$ or 2 whereas the extra degree used to compute the upper and lower bounds can take the value $p = 0$, 1, 2 or 3. As in the 1D experiments, we show the effectivity indices of the error estimates.

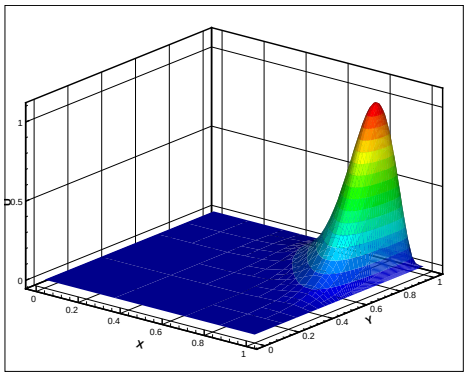

Figure 9: Representation of the exact solution $u(x)$.

The results are shown in Figs. 10 and 11 for $k = 1$ and in Figs. 12 and 13 for $k = 2$. In each case, we present the results using two methods of refinement: uniform refinement on one hand, refinement based on the energy norm on the other hand.

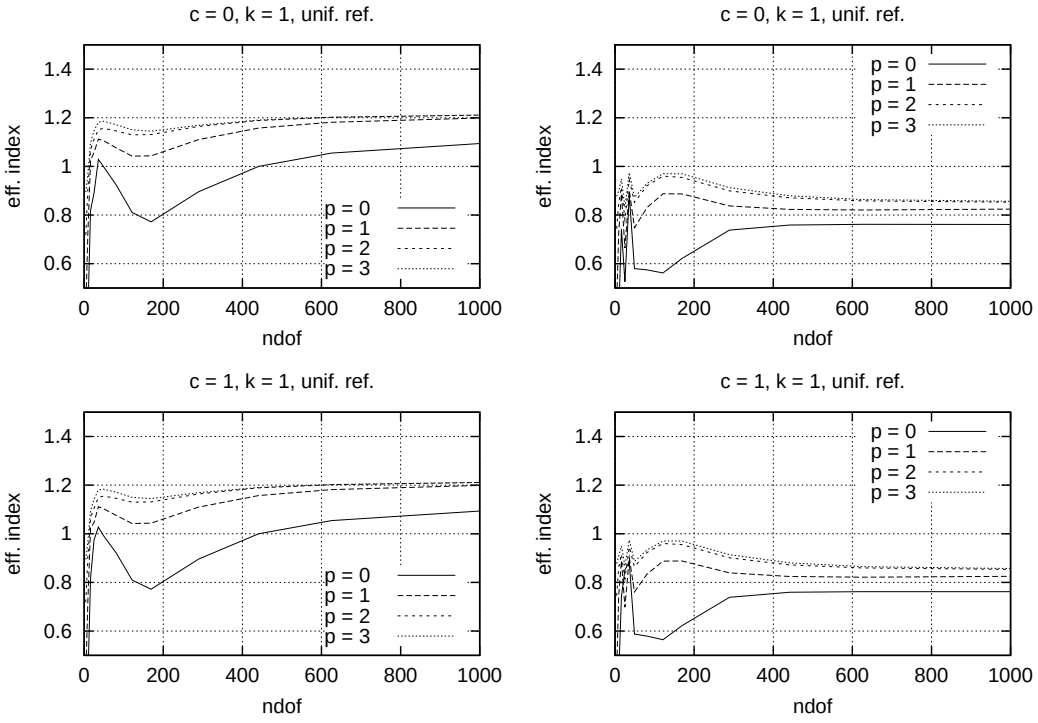

Figure 10: Global effectivity index for the upper (left) and lower (right) bounds for the case $k = 1$ and $c = 0$ (top), $c = 1$ (bottom) in the case of uniform refinement.

We remark that the results in 2D are not the same as in 1D. First, the effectivity indices are quite different for $p = 1$, 2 or 3, contrary to what we obtained in 1D. However, the effectivity indices for the upper bound tend to the same values as

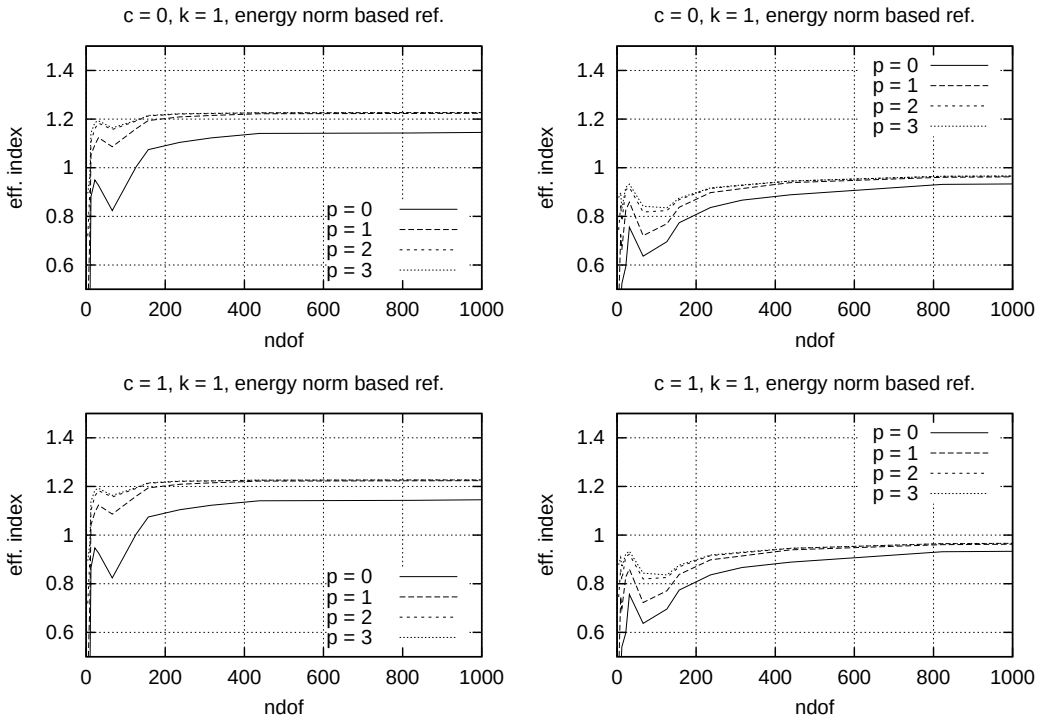

Figure 11: Global effectivity index for the upper (left) and lower (right) bounds for the case $k = 1$ and $c = 0$ (top), $c = 1$ (bottom) in the case of refinement based on the energy norm.

in 1D when p is increased, that is 1.22 for $k = 1$ and 1.07 for $k = 2$. This is an interesting point, because it suggests that these values may not depend on the geometrical dimension, and could be the same in 3D. Another interesting remark is the performance of the method regarding the lower and upper bounds when we use a refinement based on the energy norm. In that case, the local problems saturate very quickly with respect to p.

8.2 Comparison with other error estimators

In this section, we compare the SRM estimator with the Element Residual Method with (EqRM) or without (ERM) equilibrated fluxes and the Zienkiewicz and Zhu estimator (ZZ) based on a patch-recovery method of the gradient of u_h. Note that the ZZ estimator provides for an error estimate only, which in the following, is compared with the upper bound estimates obtained from the SRM, ERM, and EqRM estimators. We use again Problem (45) with $c = 0$. We choose $k = 1$ for the degree of the finite element solution and take $p = 1$ for the extra degree used to compute the SRM, ERM, and EqRM estimators. The results are shown in Fig. 14 on a sequence of uniformly refined meshes.

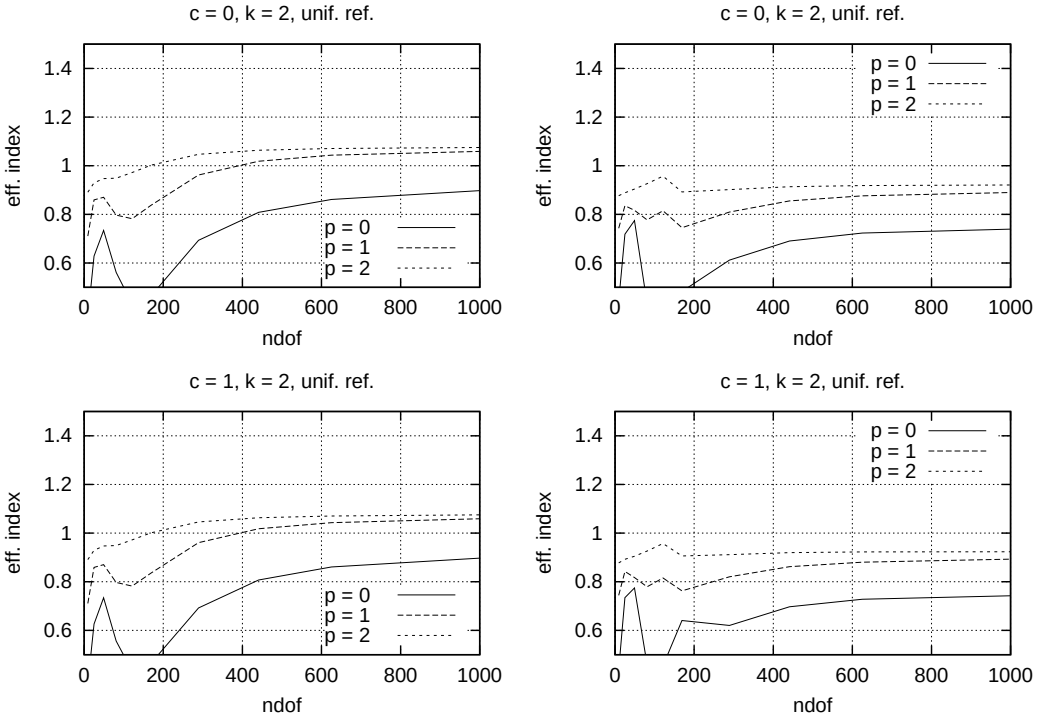

Figure 12: Global effectivity index for the upper (left) and lower (right) bounds for the case $k = 2$ and $c = 0$ (top), $c = 1$ (bottom) in the case of uniform refinement.

We observe that the SRM estimator actually provides for an upper bound on the error even with a small number of degrees of freedom. Moreover the ZZ estimator is not as accurate as the other methods, and emphasizes the idea that this estimator is not reliable for particular problems. Regarding the lower bounds, the results are comparable, although SRM seems to provide better effectivity indices with only a few degrees of freedom.

8.3 Examples of mesh adaptation

The four methods we have compared in the previous section gave different estimators of the error in the energy norm. We now study the element-wise distribution of the error given by SRM (upper bound), SRM (lower bound) and ERM, by using these error estimators in the adaptation process as described in Section 6. We again consider Problem (45) with $c = 0$, $k = 1$ for the degree of u_h, and $p = 1$ for the extra degree used in SRM and ERM. Furthermore, we choose $C_{adap} = 0.5$ to drive the adaptive process.

The adapted meshes obtained using the local contributions to the global error are shown in Fig. 15. We start with the same initial mesh of 9 uniform squared elements, and give the results after 2, 4, 6 and 8 iterations (from top to bottom).

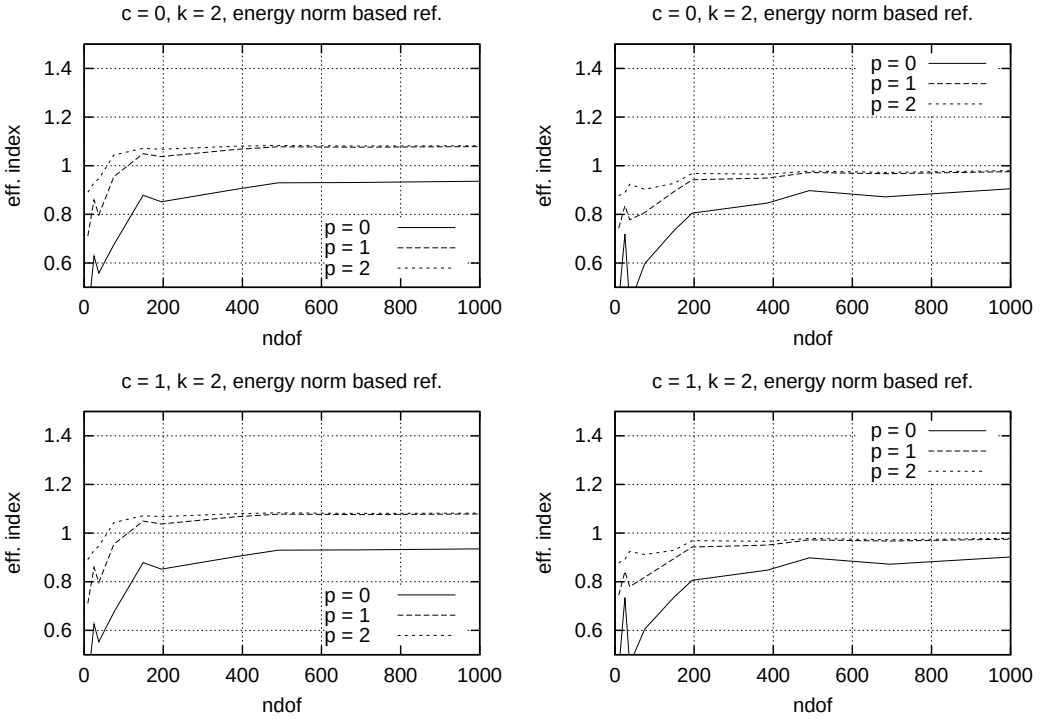

Figure 13: Global effectivity index for the upper (left) and lower (right) bounds for the case $k = 2$ and $c = 0$ (top), $c = 1$ (bottom) in the case of refinement based on the energy norm.

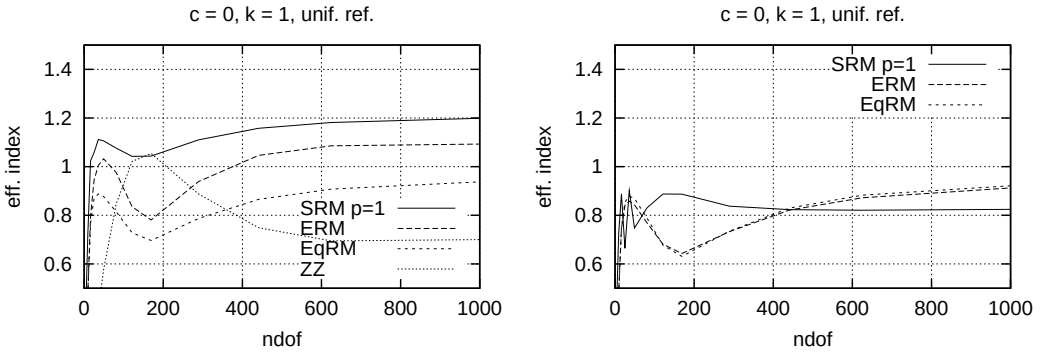

Figure 14: Global effectivity index for the upper (left) and lower (right) bounds for the SRM, ERM, EqRM, and ZZ error estimators.

We observe that the number of refined elements, after the same number of iterations, is larger when we use SRM (upper bound) than when we use the two other methods. This is due to the fact that the local contributions using SRM for the upper bound are slightly larger than those obtained by SRM for the lower bound

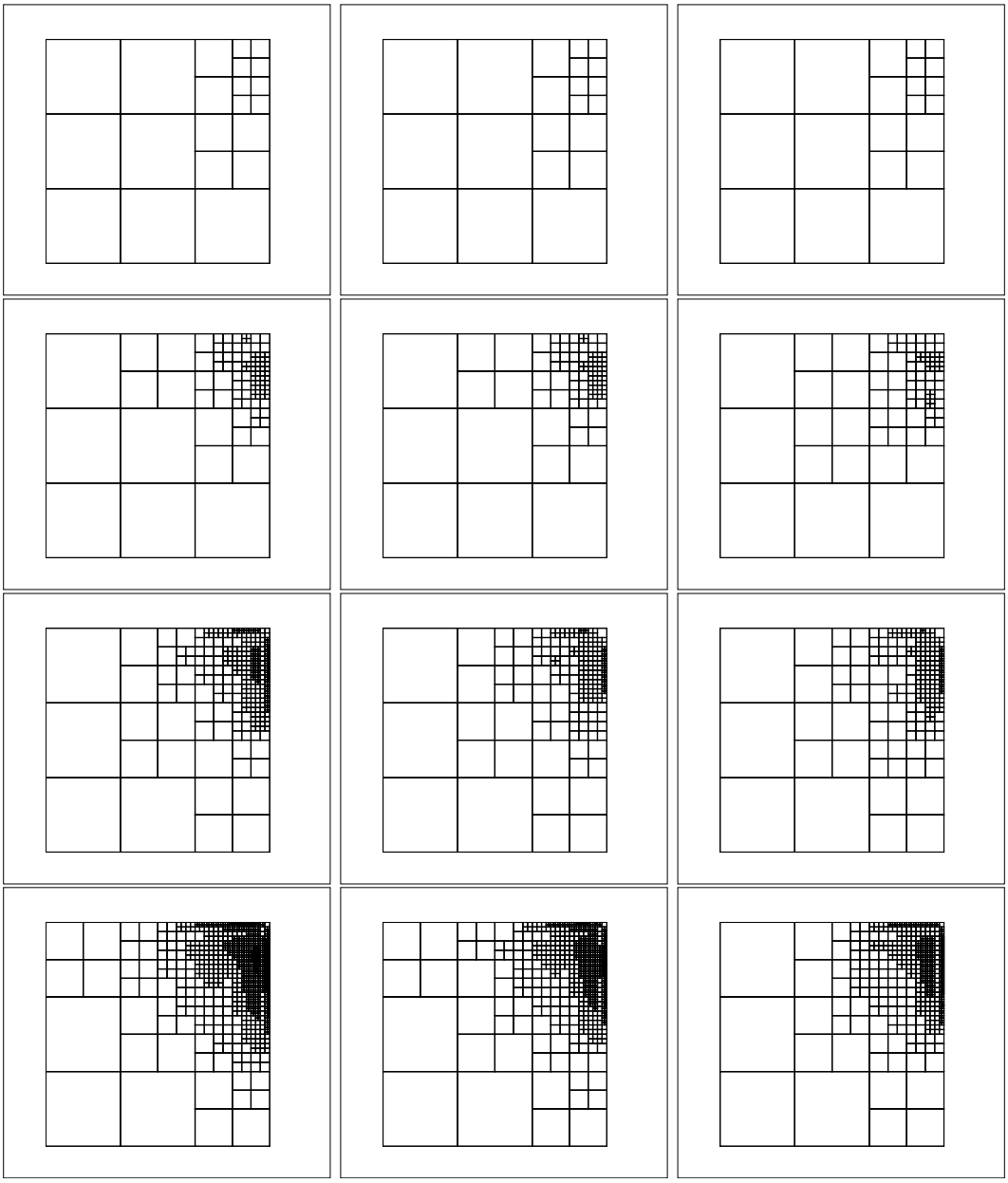

Figure 15: Adapted meshes based on the SRM upper bound estimator (left), SRM lower bound estimator (middle), and ERM estimator (right).

or ERM. In Fig. 16, we compare the convergence rate of the adaptation algorithm based on the error estimates provided either by SRM or ERM. In both cases, the upper bound estimate has been considered. We observe that for a given number of degrees of freedom, the exact error introduced by the three adaptation algorithms is comparable. Figure 16 also shows the different behavior of the methods with

respect to the number of elements refined at each iteration. For completeness, we have added the results for SRM with $p = 0$.

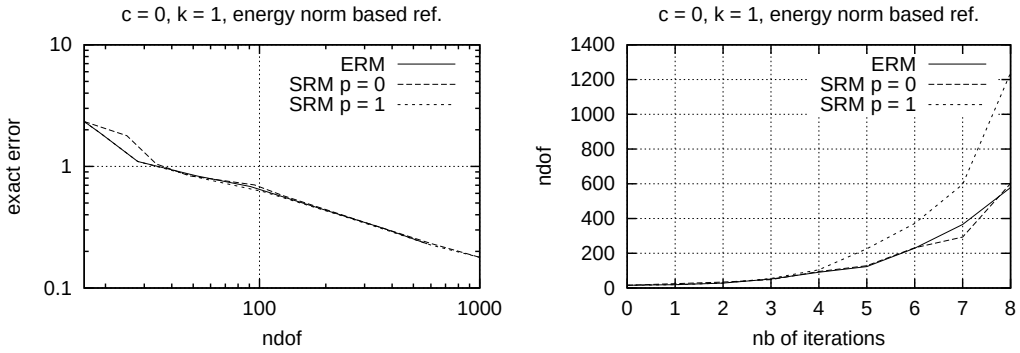

Figure 16: Graphs showing the exact error versus the number of degrees of freedom (left) and number of degrees of freedom versus the number of iterations (right), for SRM ($p = 0$ and $p = 1$) and ERM.

9 Conclusions and future work

In this work, we have investigated a subdomain-based residual method for *a posteriori* error estimation inspired from the work of Carstensen and Funken [6] and Morin, Nochetto, and Siebert [12]. However, our work is different in several aspects: first, we show that it is not necessary to compute flux jumps at the element interfaces for the calculation of the residual. Secondly, we have tested the estimator on quadrilateral elements rather than triangles and provided a complete proof of the weighted Poincaré inequality for this case (see the Appendix). Moreover, we have proposed a method to recover a lower bound on the error by a simple post-processing of the error estimator. We also show that reasonable error estimates can be obtained by using polynomial test functions of same degree as the ones used to compute the finite element solution. This is actually an important consideration as the technology can then be implemented in any existing finite element codes without having to introduce new polynomial shape functions of higher degrees, in other words, without having to drastically modify the existing data structure. Finally, the efficiency of the upper bound and lower bound estimators is demonstrated on 1D and 2D elliptic problems and compared to other existing methods (*i.e* the ZZ method, the element residual method and the equilibrated residual method). The subdomain-based residual method provides reliable and robust error estimates at reasonable cost for these test problems. However, additional numerical experiments should be performed to draw further conclusions about the quality of this error estimator for broader classes of problems. For example, it would be interesting to verify how the error estimator behaves for large values of c or more generally in the presence of boundary layers.

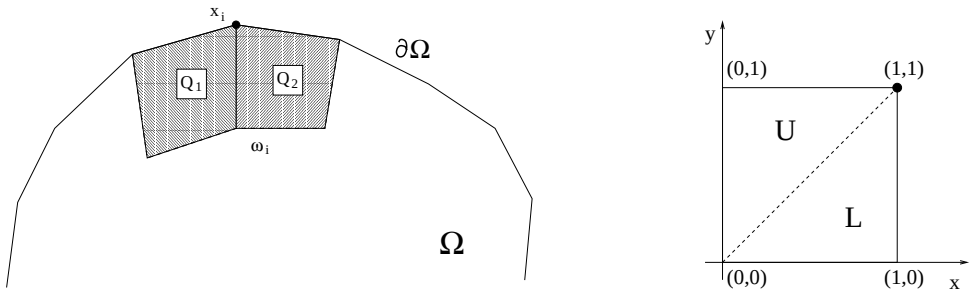

Figure 17: On the left, boundary patch; on the right, reference element $Q = [0,1] \times [0,1]$ divided into two triangles U and L.

Nevertheless, this new approach presents certain attractive features that need to be investigated in more detail. In particular, the method seems suitable for 3D applications as it appears to be much easier to implement than the equilibrated residual method, especially if the mesh contains hanging nodes. However, a question which needs to be addressed is whether the weighted Poincaré inequality holds for hexahedral or brick elements. Another issue is whether the method can be easily applied to linear elasticity and Stokes problems. Additional investigations on mesh adaptation (refinement/derefinement) are advisable as we have seen that for this method element-wise or patch-wise contributions to the error estimates can be computed. Finally, it would certainly be interesting to check how the global error estimator can be used for *a posteriori* estimation of the error in "quantities of interest".

Appendix

In this Appendix we present the proof of Lemma 2 in the case of a 2D mesh of quadrilaterals. Two different proofs are provided, one for interior patches and the other for boundary patches.

Boundary patch

Let us consider a patch ω_i on the boundary (see Fig. 17), with weighting function ϕ_i that equals one in x_i and vanishes on $\partial\omega_i \setminus \partial\Omega$, and a function f vanishing on $\partial\Omega$. Each of the two quadrilaterals Q_1 and Q_2 can be mapped onto the reference square $Q = [0,1] \times [0,1]$ by a bilinear transformation, the node x_i being mapped to $(1,1)$. The weight ϕ_i is then transformed into $\hat{\phi} = xy$. Then, by a scaling and a density argument, it is sufficient to prove the inequality

$$\int_Q f^2(x,y)\,dxdy \leq C \int_Q |\nabla f(x,y)|^2 xy\,dxdy \tag{47}$$

for any smooth function f vanishing either on $y = 1$ or $x = 1$.

Let us consider the case where $f = 0$ at $y = 1$ (the other case can be proved in a similar manner). We establish the proof of (47) in two steps. The basic idea of this proof, which is to integrate by parts along a line $y - x =$ constant, is borrowed from the paper by Carstensen *et al.* [6, Lemma 5.3].

Step 1. We first prove the inequality

$$\int_U f^2(x,y)\,dxdy \le C_1 \int_U |\nabla f(x,y)|^2 xy\,dxdy \tag{48}$$

where U is the upper triangle shown in Fig. 17.

By introducing the change of variables $x' = x$ and $y' = y - x$, which transforms the triangle $U = \{(0,0),(1,1),(0,1)\}$ into the triangle $U' = \{(0,0),(1,0),(0,1)\}$, and by integrating by parts along each line $y' =$ constant, we can write

$$\begin{aligned}\int_U f^2(x,y)\,dxdy &= \int_0^1 dy' \int_0^{1-y'} f^2(x',y')\,dx' \\ &= \int_0^1 dy' \left[f^2(x',y')x' \Big|_0^{1-y'} - \int_0^{1-y'} 2f \frac{\partial f}{\partial x'} x'\,dx' \right] \\ &= -\int_U 2f \left(\frac{\partial f}{\partial x} + \frac{\partial f}{\partial y} \right) x\,dxdy.\end{aligned}$$

The boundary term vanishes at both ends. In particular, the line $x' = 1 - y'$ corresponds to $y = 1$ and $f\big|_{y=1} = 0$. We finally have

$$\begin{aligned}\int_U f^2\,dxdy &\le \frac{1}{2}\int_U f^2\,dxdy + 2\int_U \left(\frac{\partial f}{\partial x} + \frac{\partial f}{\partial y} \right)^2 x^2\,dxdy \\ &\le \frac{1}{2}\int_U f^2\,dxdy + 4\int_U |\nabla f|^2 xy\,dxdy\end{aligned}$$

where, in the last inequality, we have used the discrete Cauchy inequality $\sum_{i=1}^n |a_i| \le \sqrt{n}\left(\sum_{i=1}^n a_i^2\right)^{1/2}$ and the fact that $x^2 \le xy$, $\forall\ (x,y) \in U$. This completes the proof of inequality (48) with a constant $C_1 = 8$.

Step 2. We now prove the inequality

$$\int_L f^2(x,y)\,dxdy \le C_2 \int_Q |\nabla f(x,y)|^2 xy\,dxdy \tag{49}$$

where L is the lower triangle shown in Fig. 17. Indeed we have

$$\begin{aligned}\int_L f^2(x,y)\,dxdy &= \int_0^1 dx \int_0^x f^2(x,y)\,dy \\ &= \int_0^1 dx \left[f^2(x,y)y \Big|_0^x - \int_0^x 2f \frac{\partial f}{\partial y} y\,dy \right] \\ &\le \int_0^1 f^2(x,x)x\,dx + \frac{1}{2}\int_L f^2\,dxdy + 2\int_L \left(\frac{\partial f}{\partial y} \right)^2 y^2\,dxdy,\end{aligned}$$

thus leading to the inequality

$$\int_L f^2(x,y)\,dxdy \leq 2\int_0^1 f^2(x,x)x\,dx + 4\int_L |\nabla f(x,y)|^2 xy\,dxdy \tag{50}$$

where, again, we have exploited the fact that $y^2 \leq xy$, $\forall\ (x,y) \in L$.

To estimate the boundary term $\int_0^1 f^2(x,x)x\,dx$ appearing in the previous inequality, we proceed as follows:

$$\begin{aligned}\int_U f^2(x,y)\,dxdy &= \int_0^1 dy \int_0^y f^2(x,y)\,dx \\ &= \int_0^1 dy \left[f^2(x,y)x\Big|_0^y - \int_0^y 2f\frac{\partial f}{\partial x}x\,dx \right] \\ &= \int_0^1 f^2(y,y)y\,dy - \int_U 2f\frac{\partial f}{\partial x}x\,dxdy.\end{aligned}$$

Hence,

$$\begin{aligned}\int_0^1 f^2(y,y)y\,dy &= \int_U f^2(x,y)\,dxdy + \int_U 2f\frac{\partial f}{\partial x}x\,dxdy \\ &\leq 2\int_U f^2(x,y)\,dxdy + \int_U |\nabla f|^2 xy\,dxdy \\ &\leq (2C_1+1)\int_U |\nabla f|^2 xy\,dxdy\end{aligned} \tag{51}$$

where, in the last inequality, we have employed relation (48). Finally, by inserting (51) in (50), we obtain (49) with $C_2 = 2(2C_1+1)$, and, by summing (48) and (49), we finally establish the statement claimed earlier.

Interior patch

Let us now consider an interior patch ω_i of quadrilaterals (see Fig. 18). In this case, the weight function ϕ_i equals one at x_i and vanishes on $\partial\omega_i$, while the function f is such that $\int_{\omega_i} f\phi_i = 0$. The patch ω_i can be mapped onto the reference patch $\hat{\omega} = [-1,1]^2$, shown in Fig. 18, by a piecewise bilinear transformation $\mathcal{F}_i : \hat{\omega} \to \omega_i$. The node x_i is mapped to $(0,0)$ and the weight ϕ_i is transformed as $\hat{\phi} = (1-|x|)(1-|y|)$. Then, by a scaling and a density argument, it is sufficient to prove the inequality

$$\left\| f - \frac{1}{\mu}\int_{\hat{\omega}} f\hat{\phi}J_i\,d\omega \right\|^2_{L^2(\hat{\omega})} \leq C\int_{\hat{\omega}} |\nabla f|^2\hat{\phi}\,d\omega, \qquad \mu = \int_{\hat{\omega}} \hat{\phi}J_i\,d\omega, \tag{52}$$

for any smooth function f, with a constant C that may depend on the regularity κ_i of the patch ω_i but is otherwise independent of h_i. We have denoted $J_i = \det(\nabla\mathcal{F}_i)$. Actually, the function $g = f - \frac{1}{\mu}\int_{\hat{\omega}} f\hat{\phi}J_i$ satisfies the condition $\int_{\omega_i} g\phi_i = \int_{\hat{\omega}} g\hat{\phi}J_i = 0$ and (52) reads also

$$\|g\|^2_{L^2(\hat{\omega})} \leq C\int_{\hat{\omega}} |\nabla g|^2\hat{\phi}\,d\omega.$$

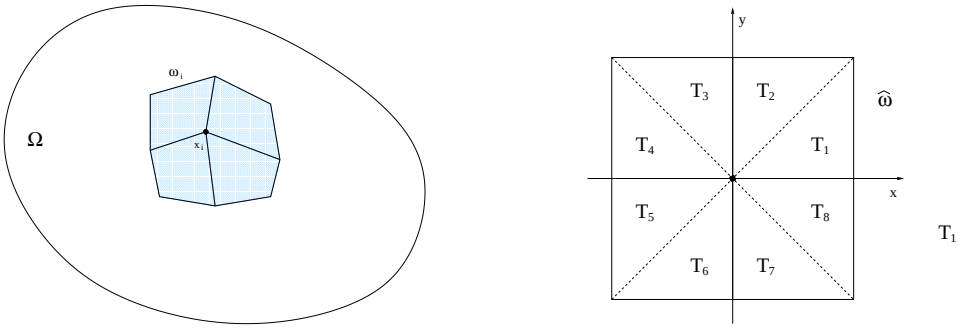

Figure 18: Interior patch of quadrilaterals (left) and patch of reference (right). Each square in the patch of reference has been divided in two triangles.

We will prove the inequalities

$$\left\| f - \frac{1}{\mu} \int_{\hat{\omega}} f \hat{\phi} J_i \, d\omega \right\|^2_{L^2(T_j)} \leq C_j \int_{\hat{\omega}} |\nabla f|^2 \hat{\phi} \, d\omega, \qquad \forall \; j = 1, \ldots, 8, \tag{53}$$

from which (52) is easily obtained by summing over the indices j. The proof is carried out only for the triangle T_1, the other cases being obtained in a similar manner.

Let us introduce the change of variables $x' = 1 - x$ and $y' = x - y$, which transforms the triangle $T_1 = \{(0,0),(1,0),(1,1)\}$ into $\{(1,0),(0,1),(0,0)\}$. Then, for any smooth function g on T_1, we have

$$\begin{aligned}
\int_{T_1} g^2(x,y)\, dxdy &= \int_0^1 dy' \int_0^{1-y'} g^2(x',y')\, dx'dy' \\
&= \int_0^1 dy' \left[g^2(x',y')x' \Big|_0^{1-y'} - \int_0^{1-y'} 2g \frac{\partial g}{\partial x'} x' \, dx' \right] \\
&= \int_0^1 g^2(1-y',y')(1-y')\, dy' + 2\int_{T_1} g \left(\frac{\partial g}{\partial x} + \frac{\partial g}{\partial y} \right) (1-x)\, dxdy \\
&\leq \int_0^1 g^2(x,0)(1-x)\, dx \\
&\qquad + \frac{1}{2} \int_{T_1} g^2 \, dxdy + 4 \int_{T_1} |\nabla g|^2 (1-x)(1-y)\, dxdy.
\end{aligned}$$

By taking $g = f - \frac{1}{\mu} \int_{\hat{\omega}} f \hat{\phi} J_i \, d\omega$, we obtain the inequality

$$\begin{aligned}
&\left\| f - \frac{1}{\mu} \int_{\hat{\omega}} f \hat{\phi} J_i \, d\omega \right\|^2_{L^2(T_1)} \\
&\qquad \leq 2 \int_0^1 \left(f - \frac{1}{\mu} \int_{\hat{\omega}} f \hat{\phi} J_i \, d\omega \right)^2 \Big|_{y=0} (1-x)\, dx + 8 \int_{T_1} |\nabla f|^2 \hat{\phi} \, d\omega
\end{aligned}$$

We now need to estimate the boundary term appearing on the right-hand side of the previous inequality.

$$\begin{aligned}
&\int_0^1 \left(f - \frac{1}{\mu} \int_{\hat{\omega}} f \hat{\phi} J_i \, d\omega \right)^2 \Big|_{y=0} (1-x)\, dx \\
&\quad = \int_0^1 \frac{1}{\mu^2} \left[\iint_{\hat{\omega}} \left(f(x,0) - f(t,s) \right) \hat{\phi}(t,s) J_i(t,s) \, dtds \right]^2 (1-x)\, dx \\
&\quad = \frac{1}{\mu^2} \int_0^1 (1-x)\, dx \\
&\qquad \left[\iint_{\hat{\omega}} \left(\int_s^0 \frac{\partial f}{\partial v}(x,v)\, dv + \int_t^x \frac{\partial f}{\partial u}(u,s)\, du \right) \hat{\phi}(t,s) J_i(t,s) \, dtds \right]^2 \\
&\quad \leq \frac{2}{\mu^2} \max_{\hat{\omega}} |J_i|^2 \left(I_1 + I_2 \right) \;\leq\; \tilde{C} \left(I_1 + I_2 \right)
\end{aligned}$$

where

$$\begin{aligned}
I_1 &= \int_0^1 (1-x) \left(\iint_{\hat{\omega}} \left| \int_s^0 \frac{\partial f}{\partial v}(x,v) \hat{\phi}(t,s)\, dv \right| dtds \right)^2 dx, \\
I_2 &= \int_0^1 (1-x) \left(\iint_{\hat{\omega}} \left| \int_t^x \frac{\partial f}{\partial u}(u,s) \hat{\phi}(t,s)\, du \right| dtds \right)^2 dx.
\end{aligned}$$

Observe that the constant $\tilde{C}$ depends only on the regularity κ_i of the patch ω_i, since $|J_i| \leq C h_i^2$ and $\mu \geq C \rho_i^2 \int_{\hat{\omega}} \hat{\phi}$.

We separately analyze the two terms I_1 and I_2. We start with I_2:

$$\begin{aligned}
I_2 &= \int_0^1 (1-x) dx \left(\int_{-1}^1 dt \int_{-1}^1 \left| \int_t^x \frac{\partial f}{\partial u}(u,s) \sqrt{\hat{\phi}(u,s)} \frac{\hat{\phi}(t,s)}{\sqrt{\hat{\phi}(u,s)}}\, du \right| ds \right)^2 \\
&\leq \int_0^1 (1-x) dx \\
&\qquad \left[\int_{-1}^1 \left(\iint_{\hat{\omega}} \left| \frac{\partial f}{\partial u}(u,s) \right|^2 \hat{\phi}(u,s)\, duds \right)^{\frac{1}{2}} \left(\left| \int_{-1}^1 \int_t^x \frac{\hat{\phi}^2(t,s)}{\hat{\phi}(u,s)}\, duds \right| \right)^{\frac{1}{2}} dt \right]^2
\end{aligned}$$

that is,

$$I_2 \leq \left(\iint\limits_{\hat{\omega}} |\nabla f|^2 \hat{\phi}\, d\omega \right) \int_0^1 (1-x) dx \left[\int_{-1}^1 dt \left(\int_{-1}^1 ds \left| \int_t^x \frac{(1-|t|)^2(1-|s|)}{1-|u|}\, du \right| \right)^{\frac{1}{2}} \right]^2 \leq C_1 \int_{\hat{\omega}} |\nabla f|^2 \hat{\phi}\, d\omega$$

where the constant C_1 is given by

$$C_1 = \int_0^1 (1-x)\, dx \left[\int_{-1}^1 (1-|t|)\, dt \left(\int_{-1}^1 (1-|s|)\, ds \left| \int_t^x \frac{1}{1-|u|}\, du \right| \right)^{\frac{1}{2}} \right]^2 \leq \int_0^1 (1-x)\, dx \left[\int_{-1}^1 (1-|t|) \left[-\log(1-|x|) - \log(1-|t|) \right]^{\frac{1}{2}}\, dt \right]^2 < +\infty.$$

We now consider the term I_1:

$$I_1 = \int_0^1 dx \left(\int_{-1}^1 dt \int_{-1}^1 ds \left| \int_s^0 \frac{\partial f}{\partial v}(x,v) \sqrt{\hat{\phi}(x,v)} \frac{\hat{\phi}(t,s)\sqrt{1-x}}{\sqrt{\hat{\phi}(x,v)}}\, dv \right| \right)^2 \leq \int_0^1 dx \left[\int_{-1}^1 dt \int_{-1}^1 ds \left(\int_{-1}^1 \left| \frac{\partial f}{\partial v}(x,v) \right|^2 \hat{\phi}(x,v)\, dv \right)^{\frac{1}{2}} \left(\left| \int_s^0 \frac{\hat{\phi}^2(t,s)(1-|x|)}{\hat{\phi}(x,v)}\, dv \right| \right)^{\frac{1}{2}} \right]^2 \leq \int_0^1 dx \left(\int_{-1}^1 \left| \frac{\partial f}{\partial v}(x,v) \right|^2 \hat{\phi}(x,v)\, dv \right) \left[\int_{-1}^1 dt \int_{-1}^1 ds \left(\left| \int_s^0 \frac{(1-|t|)^2(1-|s|)^2}{(1-|v|)}\, dv \right| \right)^{\frac{1}{2}} \right]^2 \leq C_2 \int_{\hat{\omega}} |\nabla f|^2 \hat{\phi}\, d\omega$$

where the constant C_2 is given by

$$C_2 = \left[\int_{-1}^1 (1-|t|)\, dt \int_{-1}^1 (1-|s|)\, ds \left(\left| \int_s^0 \frac{1}{1-|v|}\, dv \right| \right)^{\frac{1}{2}} \right]^2 = \left(\int_{-1}^1 (1-|s|) \left[-\log(1-|s|) \right]^{\frac{1}{2}}\, ds \right)^2 < +\infty$$

Observe that the constant C_2 is bounded thanks to the presence of the weight $(1-x)$.

We finally obtain

$$\int_0^1 \left(f - \frac{1}{\mu} \int_{\hat{\omega}} f \hat{\phi} \, d\omega \right)^2 \Big|_{y=0} (1-x) \, dx \leq \tilde{C}(C_1 + C_2) \int_{\hat{\omega}} |\nabla f|^2 \hat{\phi} \, d\omega.$$

Then, inequality (53) is satisfied on the triangle T_1. An analogous proof holds for the other triangles, thus leading to the final result.

Acknowledgement. The support of this work through ONR grant N00014-95-0401 is gratefully acknowledged. The authors wish also to acknowledge Prof. Babuška for the fruitful discussion on the subject.

References

[1] M. Ainsworth and J.T. Oden, *A Posteriori Error Estimation in Finite Element Analysis,* John Wiley and Sons, 2000.

[2] M. Ainsworth and J.T. Oden, A unified approach to a *a posteriori* error estimation based on element residual methods, *Numer. Math.,* **65**, 23 (1993).

[3] I. Babuška and W.C. Rheinboldt, *A posteriori* error estimates for the finite element method, *Internat. J. Numer. Methods Engrg.,* **12**, 1597 (1978).

[4] I. Babuška, T. Strouboulis and S.K. Gangaraj, Guaranteed computable bounds for the exact error in the finite element solution. Part I: One-dimensional model problem, *Comput. Methods Appl. Mech. Engrg.,* **176**, 51 (1999).

[5] R.E. Bank and A. Weiser, Some *a posteriori* error estimators for elliptic partial differential equations, *Math. Comp.,* **44**, 283 (1985).

[6] C. Carstensen and S.A. Funken, Fully reliable localized error control in the FEM, *SIAM Journal of Scientific Computing,* **21**, 1465 (2000).

[7] P.G. Ciarlet, *The Finite Element Method for Elliptic Problems,* North-Holland, 1978.

[8] D.K. Datta, *Computer Analysis of Error Estimation in Finite Element Computations for Elliptic and Parabolic Problems,* Ph.D. Dissertation, Texas A&M University (2001).

[9] L. Demkowicz, J.T. Oden, and T. Strouboulis, Adaptative finite elements for flow problems with moving boundaries. Part 1: Variational principles and *a posteriori* error estimates, *Comput. Methods Appl. Mech. Engrg.,* **46**, 217 (1984).

[10] A. Ern and J.-L. Guermond, *Eléments Finis: Théorie, Applications, Mise en Oeuvre,* Springer-Verlag, Berlin, 2002.

[11] P. Ladevèze and D. Leguillon, Error estimate procedure in the finite element method and applications, *SIAM J. Numer. Anal.,* **20**, 485 (1983).

[12] P. Morin, R.H. Nochetto, and K.G. Siebert, Local problems on stars: *a posteriori* error estimators, convergence, and performance, *Math. Comp.,* (to appear).

[13] T. Strouboulis, I. Babuška, and S.K. Gangaraj, Guaranteed computable bounds for the exact error in the finite element solution. Part II: bounds for the energy norm of the error in two dimensions, *Internat. J. Numer. Methods Engrg.,* **47**, 427 (2000).

[14] O.C. Zienkiewicz and J.Z. Zhu, The superconvergent patch recovery and *a posteriori* error estimates. Part 1: The recovery technique, *Internat. J. Numer. Methods Engrg.,* **33**, 1331 (1992).

[15] O.C. Zienkiewicz and J.Z. Zhu, The superconvergent patch recovery and *a posteriori* error estimates. Part 2: Error estimates and adaptivity, *Internat. J. Numer. Methods Engrg.,* **33**, 1365 (1992).